

0550  
01/02

OIFE

## RAW SEQUENCE LISTING

DATE: 01/03/2002

PATENT APPLICATION: US/09/802,472A

TIME: 15:07:04

Input Set : A:\EINAT4.1C.txt

Output Set: N:\CRF3\01032002\I802472A.raw

P.S

3 <110> APPLICANT: EINAT, Paz  
 4 SKALITER, Rami  
 5 FEINSTEIN, Elena  
 7 <120> TITLE OF INVENTION: SEQUENCES CHARACTERISTIC OF HYPOXIA-REGULATED GENE  
 TRANSCRIPTION

9 <130> FILE REFERENCE: EINAT=4.1C  
 11 <140> CURRENT APPLICATION NUMBER: US/09/802,472A  
 11 <141> CURRENT FILING DATE: 2001-03-09  
 11 <150> PRIOR APPLICATION NUMBER: US 09/383,096  
 12 <151> PRIOR FILING DATE: 1999-08-27  
 14 <150> PRIOR APPLICATION NUMBER: US 09/138,109  
 15 <151> PRIOR FILING DATE: 1998-08-21  
 17 <150> PRIOR APPLICATION NUMBER: US 60/098,158  
 18 <151> PRIOR FILING DATE: 1998-08-27  
 20 <150> PRIOR APPLICATION NUMBER: US 60/132,684  
 21 <151> PRIOR FILING DATE: 1999-05-05  
 23 <160> NUMBER OF SEQ ID NOS: 21  
 25 <170> SOFTWARE: PatentIn version 3.0  
 26 <210> SEQ ID NO: 1  
 28 <211> LENGTH: 1655  
 29 <212> TYPE: DNA  
 30 <213> ORGANISM: Homo sapiens  
 32 <220> FEATURE:  
 33 <221> NAME/KEY: CDS  
 34 <222> LOCATION: (265)..(1575)  
 36 <220> FEATURE:  
 37 <221> NAME/KEY: misc\_feature  
 38 <222> LOCATION: (5)..(5)  
 39 <223> OTHER INFORMATION: "n" is unknown  
 42 <220> FEATURE:  
 43 <221> NAME/KEY: misc\_feature  
 44 <222> LOCATION: (22)..(22)  
 45 <223> OTHER INFORMATION: "n" is unknown  
 48 <220> FEATURE:  
 49 <221> NAME/KEY: misc\_feature  
 50 <222> LOCATION: (41)..(41)  
 51 <223> OTHER INFORMATION: "n" is unknown  
 54 <220> FEATURE:  
 55 <221> NAME/KEY: misc\_feature  
 56 <222> LOCATION: (47)..(47)  
 57 <223> OTHER INFORMATION: "n" is unknown  
 60 <400> SEQUENCE: 1

ENTERED

w-OK 61 gcacnaggtg tgtggcagca anagccgccca gttcggggacc nccgcanctg ggggtggcaac 60  
 63 ggcgcaggag gggtcgcggg gagggagtgg tgagcgcagg cggcaggggt ctgggaaaga 120  
 65 cgaagtcgct atttgctgtc tgagcgcgct cgcagctcct ggaagtgttg ccgcctctcg 180  
 67 gtttcgctct cgctcgctgc gctcctagaa ggggcggccg cctccaggac tgaccagggc 240  
 69 caagtggcgc tcggcgggca ctac atg gcg gag ggt gaa ggg tac ttc gcc 291  
 70 Met Ala Glu Gly Glu Gly Tyr Phe Ala

DATE: 01/03/2002

PATENT APPLICATION: US/09/802,472A

TIME: 15:07:04

Input Set : A:\EINAT4.1C.txt

Output Set: N:\CRF3\01032002\I802472A.raw

71	1										5										
73	atg	tct	gag	gac	gag	ctg	gcc	tgc	agc	ccc	tac	atc	ccc	cta	ggc	ggc	339				
74	Met	Ser	Glu	Asp	Glu	Leu	Ala	Cys	Ser	Pro	Tyr	Ile	Pro	Leu	Gly	Gly					
75	10					15					20				25						
77	gac	ttc	ggc	ggc	ggc	gac	ttc	ggc	ggc	ggc	gac	ttc	ggc	ggt	ggc	ggc	387				
78	Asp	Phe	Gly	Gly	Gly	Asp	Phe	Gly	Gly	Gly	Asp	Phe	Gly	Gly	Gly	Gly					
79					30					35					40						
81	agc	ttc	ggt	ggg	cat	tgc	ttg	gac	tat	tgc	gaa	agc	cct	acg	gcg	cac	435				
82	Ser	Phe	Gly	Gly	His	Cys	Leu	Asp	Tyr	Cys	Glu	Ser	Pro	Thr	Ala	His					
83				45					50					55							
85	tgc	aat	gtg	ctg	aac	tgg	gag	caa	gtg	cag	cgg	ctg	gac	ggc	atc	ctg	483				
86	Cys	Asn	Val	Leu	Asn	Trp	Glu	Gln	Val	Gln	Arg	Leu	Asp	Gly	Ile	Leu					
87		60					65					70									
89	agc	gag	acc	att	ccg	att	cac	ggg	cgc	ggc	aac	ttc	ccc	acg	ctc	gag	531				
90	Ser	Glu	Thr	Ile	Pro	Ile	His	Gly	Arg	Gly	Asn	Phe	Pro	Thr	Leu	Glu					
91	75					80					85										
93	ctg	cag	ccg	agc	ctg	atc	gtg	aag	gtg	gtg	cgg	cgg	cgc	ctg	gcc	gag	579				
94	Leu	Gln	Pro	Ser	Leu	Ile	Val	Lys	Val	Val	Arg	Arg	Arg	Leu	Ala	Glu					
95	90				95						100				105						
97	aag	cgc	att	ggc	gtc	cgc	gac	gtg	cgc	ctc	aac	ggc	tcg	gca	gcc	agc	627				
98	Lys	Arg	Ile	Gly	Val	Arg	Asp	Val	Arg	Leu	Asn	Gly	Ser	Ala	Ala	Ser					
99				110					115					120							
101	cat	gtc	ctg	cac	cag	gac	agc	ggc	ctg	ggc	tac	aag	gac	ctg	gac	ctc	675				
102	His	Val	Leu	His	Gln	Asp	Ser	Gly	Leu	Gly	Tyr	Lys	Asp	Leu	Asp	Leu					
103				125					130					135							
105	atc	ttc	tgc	gcc	gac	ctg	cgc	ggg	gaa	ggg	gag	ttt	cag	act	gtg	aag	723				
106	Ile	Phe	Cys	Ala	Asp	Leu	Arg	Gly	Glu	Gly	Glu	Phe	Gln	Thr	Val	Lys					
107			140					145				150									
109	gac	gtc	gtg	ctg	gac	tgc	ctg	ttg	gac	ttc	tta	ccc	gag	ggg	gtg	aac	771				
110	Asp	Val	Val	Leu	Asp	Cys	Leu	Leu	Asp	Phe	Leu	Pro	Glu	Gly	Val	Asn					
111		155				160					165										
113	aaa	gag	aag	atc	aca	cca	ctc	acg	ctc	aag	gaa	gct	tat	gtg	cag	aaa	819				
114	Lys	Glu	Lys	Ile	Thr	Pro	Leu	Thr	Leu	Lys	Glu	Ala	Tyr	Val	Gln	Lys					
115	170				175						180				185						
117	atg	gtt	aaa	gtg	tgc	aat	gac	tct	gac	cga	tgg	agt	ctt	ata	tcc	ctg	867				
118	Met	Val	Lys	Val	Cys	Asn	Asp	Ser	Asp	Arg	Trp	Ser	Leu	Ile	Ser	Leu					
119				190					195					200</							

## RAW SEQUENCE LISTING

DATE: 01/03/2002

PATENT APPLICATION: US/09/802,472A

TIME: 15:07:04

Input Set : A:\EINAT4.1C.txt

Output Set: N:\CRF3\01032002\I802472A.raw

```

137 gcc ttt gat cac ctt tgt aac aag atc att gcc acc agg aac cca gag      1107
138 Ala Phe Asp His Leu Cys Asn Lys Ile Ile Ala Thr Arg Asn Pro Glu
139                               270                               275                               280
141 gaa atc cga ggg gga ggc ctg ctt aag tac tgc aac ctc ttg gtg agg      1155
142 Glu Ile Arg Gly Gly Gly Leu Leu Lys Tyr Cys Asn Leu Leu Val Arg
143                               285                               290                               295
145 ggc ttt agg ccc gcc tct gat gaa atc aag acc ctt caa agg tat atg      1203
146 Gly Phe Arg Pro Ala Ser Asp Glu Ile Lys Thr Leu Gln Arg Tyr Met
147                               300                               305                               310
149 tgt tcc agg ttt ttc atc gac ttc tca gac att gga gag cag cag aga      1251
150 Cys Ser Arg Phe Phe Ile Asp Phe Ser Asp Ile Gly Glu Gln Gln Arg
151                               315                               320                               325
153 aaa ctg gag tcc tat ttg cag aac ctc ttt gtg gga ttg gaa gcc cgc      1299
154 Lys Leu Glu Ser Tyr Leu Gln Asn Leu Phe Val Gly Leu Glu Ala Arg
155 330                               335                               340                               345
157 aag tat gag tat ctc atg acc ctt cat gga gtg gta aat gag agc tca      1347
158 Lys Tyr Glu Tyr Leu Met Thr Leu His Gly Val Val Asn Glu Ser Ser
159                               350                               355                               360
161 gtg tgc ctg atg gga cat gaa aga aga cag act tta aac ctt atc acc      1395
162 Val Cys Leu Met Gly His Glu Arg Arg Gln Thr Leu Asn Leu Ile Thr
163                               365                               370                               375
165 atg ctg gct atc cgg gtg tta gct gac caa aat gtc att cct aat gtg      1443
166 Met Leu Ala Ile Arg Val Leu Ala Asp Gln Asn Val Ile Pro Asn Val
167                               380                               385                               390
169 gct aat gtc act tgc tat tac cag cca gcc ccc tat gta gca gat gcc      1491
170 Ala Asn Val Thr Cys Tyr Tyr Gln Pro Ala Pro Tyr Val Ala Asp Ala
171                               395                               400                               405
173 aac ttt agc aat tac tac att gca cag gtt cag cca gta ttc acg tgc      1539
174 Asn Phe Ser Asn Tyr Tyr Ile Ala Gln Val Gln Pro Val Phe Thr Cys
175 410                               415                               420                               425
177 cag caa cag acc tac tcc act tgg cta ccc tgc aat taagaatcat      1585
178 Gln Gln Gln Thr Tyr Ser Thr Trp Leu Pro Cys Asn
179                               430                               435
181 ttaaaaaatgt cctgtgggga agccatttca gacaagacag gagagaaaaa aaaaaaaaaa      1645
183 aaaaaaaaaa      1655
186 <210> SEQ ID NO: 2
187 <211> LENGTH: 437
188 <212> TYPE: PRT
189 <213> ORGANISM: Homo sapiens
191 <400> SEQUENCE: 2
193 Met Ala Glu Gly Glu Gly Tyr Phe Ala Met Ser Glu Asp Glu Leu Ala
194 1                               5                               10                               15
197 Cys Ser Pro Tyr Ile Pro Leu Gly Gly Asp Phe Gly Gly Gly Asp Phe
198                               20                               25                               30
201 Gly Gly Gly Asp Phe Gly Gly Gly Ser Phe Gly Gly His Cys Leu
202                               35                               40                               45
205 Asp Tyr Cys Glu Ser Pro Thr Ala His Cys Asn Val Leu Asn Trp Glu
206                               50                               55                               60
209 Gln Val Gln Arg Leu Asp Gly Ile Leu Ser Glu Thr Ile Pro Ile His

```

## RAW SEQUENCE LISTING

DATE: 01/03/2002

PATENT APPLICATION: US/09/802,472A

TIME: 15:07:04

Input Set : A:\EINAT4.1C.txt

Output Set: N:\CRF3\01032002\I802472A.raw

```

210 65          70          75          80
213 Gly Arg Gly Asn Phe Pro Thr Leu Glu Leu Gln Pro Ser Leu Ile Val
214          85          90          95
217 Lys Val Val Arg Arg Arg Leu Ala Glu Lys Arg Ile Gly Val Arg Asp
218          100          105          110
221 Val Arg Leu Asn Gly Ser Ala Ala Ser His Val Leu His Gln Asp Ser
222          115          120          125
225 Gly Leu Gly Tyr Lys Asp Leu Asp Leu Ile Phe Cys Ala Asp Leu Arg
226          130          135          140
229 Gly Glu Gly Glu Phe Gln Thr Val Lys Asp Val Val Leu Asp Cys Leu
230 145          150          155          160
233 Leu Asp Phe Leu Pro Glu Gly Val Asn Lys Glu Lys Ile Thr Pro Leu
234          165          170          175
237 Thr Leu Lys Glu Ala Tyr Val Gln Lys Met Val Lys Val Cys Asn Asp
238          180          185          190
241 Ser Asp Arg Trp Ser Leu Ile Ser Leu Ser Asn Asn Ser Gly Lys Asn
242          195          200          205
245 Val Glu Leu Lys Phe Val Asp Ser Leu Arg Arg Gln Phe Glu Phe Ser
246          210          215          220
249 Val Asp Ser Phe Gln Ile Lys Leu Asp Ser Leu Leu Leu Phe Tyr Glu
250 225          230          235          240
253 Cys Ser Glu Asn Pro Met Thr Glu Thr Phe His Pro Thr Ile Ile Gly
254          245          250          255
257 Glu Ser Val Tyr Gly Asp Phe Gln Glu Ala Phe Asp His Leu Cys Asn
258          260          265          270
261 Lys Ile Ile Ala Thr Arg Asn Pro Glu Glu Ile Arg Gly Gly Gly Leu
262          275          280          285
265 Leu Lys Tyr Cys Asn Leu Leu Val Arg Gly Phe Arg Pro Ala Ser Asp
266          290          295          300
269 Glu Ile Lys Thr Leu Gln Arg Tyr Met Cys Ser Arg Phe Phe Ile Asp
270 305          310          315          320
273 Phe Ser Asp Ile Gly Glu Gln Gln Arg Lys Leu Glu Ser Tyr Leu Gln
274          325          330          335
277 Asn Leu Phe Val Gly Leu Glu Ala Arg Lys Tyr Glu Tyr Leu Met Thr
278          340          345          350
281 Leu His Gly Val Val Asn Glu Ser Ser Val Cys Leu Met Gly His Glu
282          355          360          365
285 Arg Arg Gln Thr Leu Asn Leu Ile Thr Met Leu Ala Ile Arg Val Leu
286          370          375          380
289 Ala Asp Gln Asn Val Ile Pro Asn Val Ala Asn Val Thr Cys Tyr Tyr
290 385          390          395          400
293 Gln Pro Ala Pro Tyr Val Ala Asp Ala Asn Phe Ser Asn Tyr Tyr Ile
294          405          410          415
297 Ala Gln Val Gln Pro Val Phe Thr Cys Gln Gln Gln Thr Tyr Ser Thr
298          420          425          430
301 Trp Leu Pro Cys Asn
302          435
305 <210> SEQ ID NO: 3
306 <211> LENGTH: 3454

```

## RAW SEQUENCE LISTING

DATE: 01/03/2002

PATENT APPLICATION: US/09/802,472A

TIME: 15:07:04

Input Set : A:\EINAT4.1C.txt

Output Set: N:\CRF3\01032002\I802472A.raw

```

307 <212> TYPE: DNA
308 <213> ORGANISM: Homo sapiens
310 <220> FEATURE:
311 <221> NAME/KEY: CDS
312 <222> LOCATION: (323)..(1762)
314 <220> FEATURE:
315 <221> NAME/KEY: misc_feature
316 <222> LOCATION: (606)..(606)
317 <223> OTHER INFORMATION: "n" is unknown
320 <220> FEATURE:
321 <221> NAME/KEY: misc_feature
322 <222> LOCATION: (2561)..(2561)
323 <223> OTHER INFORMATION: "n" is unknown
326 <220> FEATURE:
327 <221> NAME/KEY: misc_feature
328 <222> LOCATION: (2594)..(2504)
329 <223> OTHER INFORMATION: "n" is unknown
332 <220> FEATURE:
333 <221> NAME/KEY: misc_feature
334 <222> LOCATION: (2613)..(2613)
335 <223> OTHER INFORMATION: "n" is unknown
338 <400> SEQUENCE: 3
339 ctccgcggcg gggatgctga ggagcgtgg gtccgggagc agccctggcc cctgcggact      60
341 tccgagggcg tgaaaacccc tgcgctgcgg cccttcccag gccccgagg ccgttcgccc      120
343 ttcccgaagc ccgactgggg gaagagtcca gcaccaaagc ggccgttctc ggattccgga      180
345 gcgttctgga gccccgagag acgccccggg gttctagaag ctccccggcg gcgcccagtc      240
347 cgggttcat tcgggcgtcc ctccgaaacc cactcgggtg cacgggtcgt cggcgagccg      300
349 cgaccgggtc ctggcgcgca cc atg atc gtg gcg gac tcc gag tgc cgc gca      352
350                               Met Ile Val Ala Asp Ser Glu Cys Arg Ala
351                               1             5             10
353 gag ctc aag gac tac ctg cgg ttc gcc ccg ggc ggc gtc ggc gac tcg      400
354 Glu Leu Lys Asp Tyr Leu Arg Phe Ala Pro Gly Gly Val Gly Asp Ser
355                               15             20             25
357 ggc ccc gga gag gag cag agg gag agc cgg gct cgg cga ggc cct cga      448
358 Gly Pro Gly Glu Glu Gln Arg Glu Ser Arg Ala Arg Arg Gly Pro Arg
359                               30             35             40
361 ggg ccc agc gcc ttc atc ccc gtg gag gag gtc ctt cgg gag ggg gct      496
362 Gly Pro Ser Ala Phe Ile Pro Val Glu Glu Val Leu Arg Glu Gly Ala
363                               45             50             55
365 gag agc ctc gag cag cac ctg ggg ctg gag gca ctg atg tcc tct ggg      544
366 Glu Ser Leu Glu Gln His Leu Gly Leu Glu Ala Leu Met Ser Ser Gly
367                               60             65             70
369 cga gta gac aac ctg gca gtg gtg atg ggc ctg cac cct gac tac ttt      592
370 Arg Val Asp Asn Leu Ala Val Val Met Gly Leu His Pro Asp Tyr Phe
371 75                               80             85             90
w-OK 373 acc agc ttc tgg cnc ctg cac tac ctg ctg ctg cac acg gat ggt ccc      640
w--> 374 Thr Ser Phe Trp Xaa Leu His Tyr Leu Leu His Thr Asp Gly Pro
375                               95             100            105
377 ttg gcc agc tcc tgg cgc cac tac att gcc atc atg gct gcc gcc cgc      688

```

Use of n and/or Xaa has been detected in the Sequence Listing.  
 Review the Sequence Listing to insure a corresponding  
 explanation is presented in the <220> to <223> fields of  
 each sequence using n or Xaa.

## VERIFICATION SUMMARY

DATE: 01/03/2002

PATENT APPLICATION: US/09/802,472A

TIME: 15:07:05

Input Set : A:\EINAT4.1C.txt

Output Set: N:\CRF3\01032002\I802472A.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No  
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:61 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:373 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:374 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:497 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:499 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:562 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
L:873 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
L:875 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
L:911 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
L:913 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
L:915 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
L:952 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:2618 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:20  
L:2631 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20